1/10

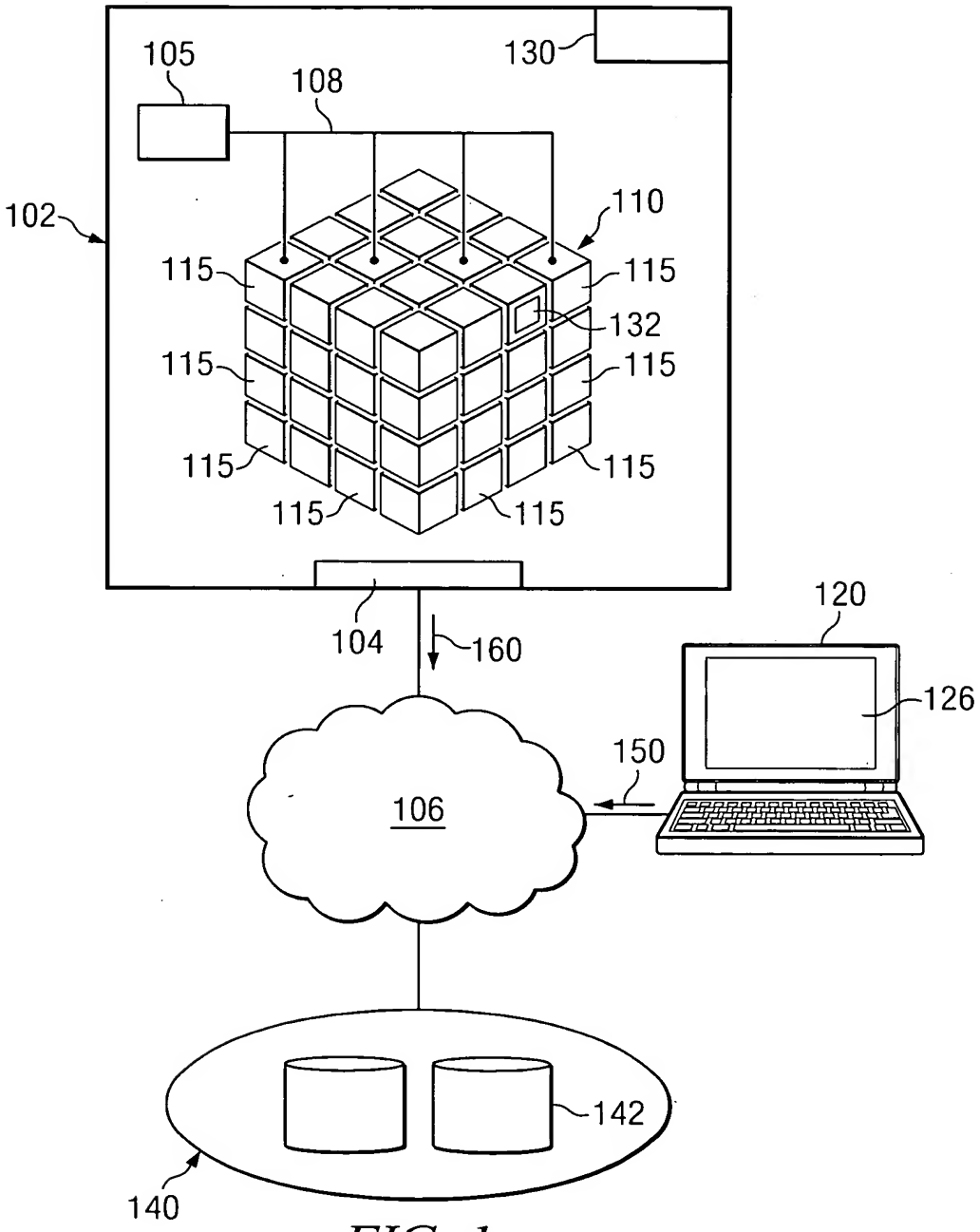

105
108
130
102
110
115
115
132
115
115
115
115
115
115
104
160
120
126
150
106
142
140

FIG. 1

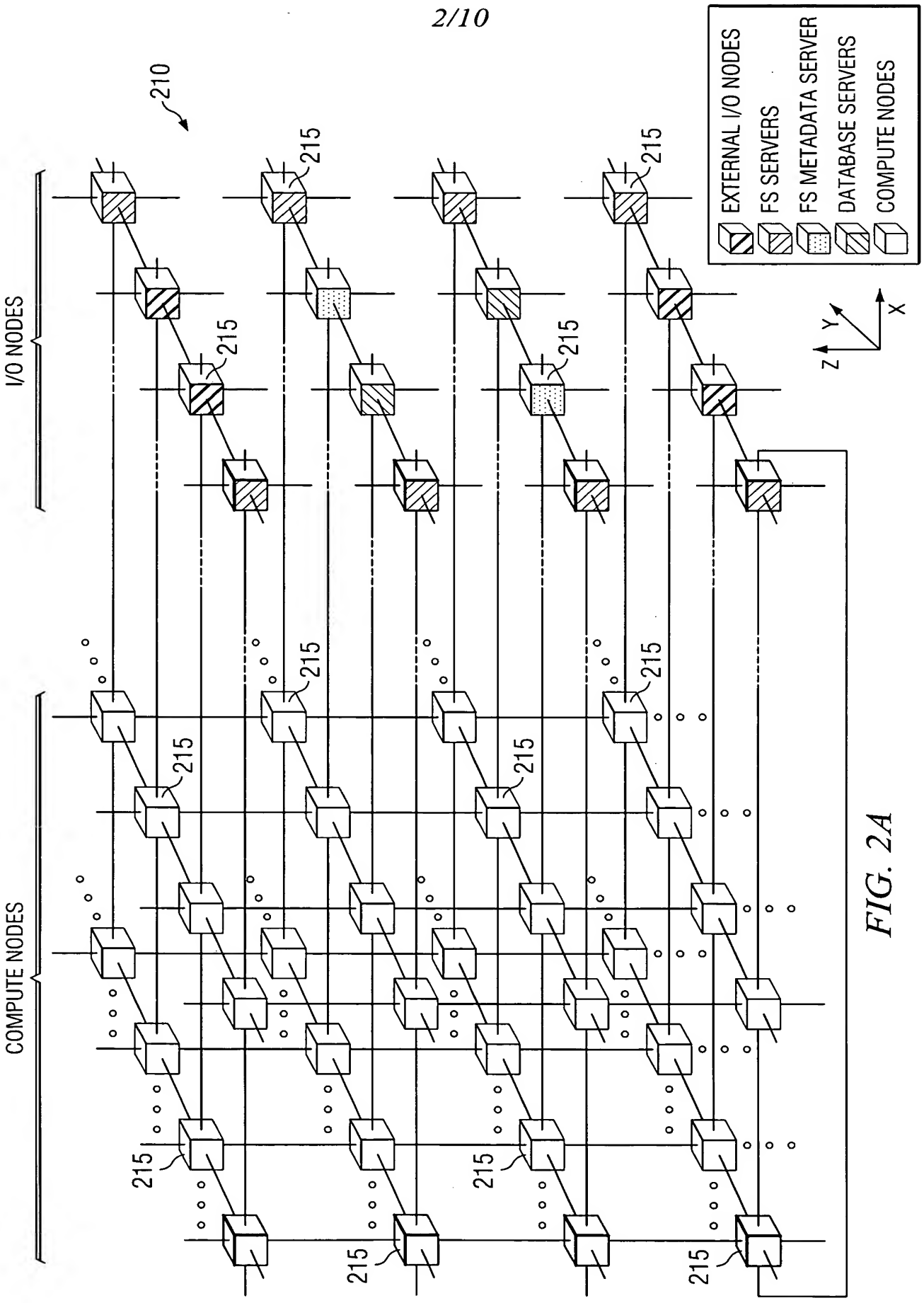

210
215
I/O NODES
COMPUTE NODES
EXTERNAL I/O NODES
FS SERVERS
FS METADATA SERVER
DATABASE SERVERS
COMPUTE NODES
X
Y
Z

FIG. 2A

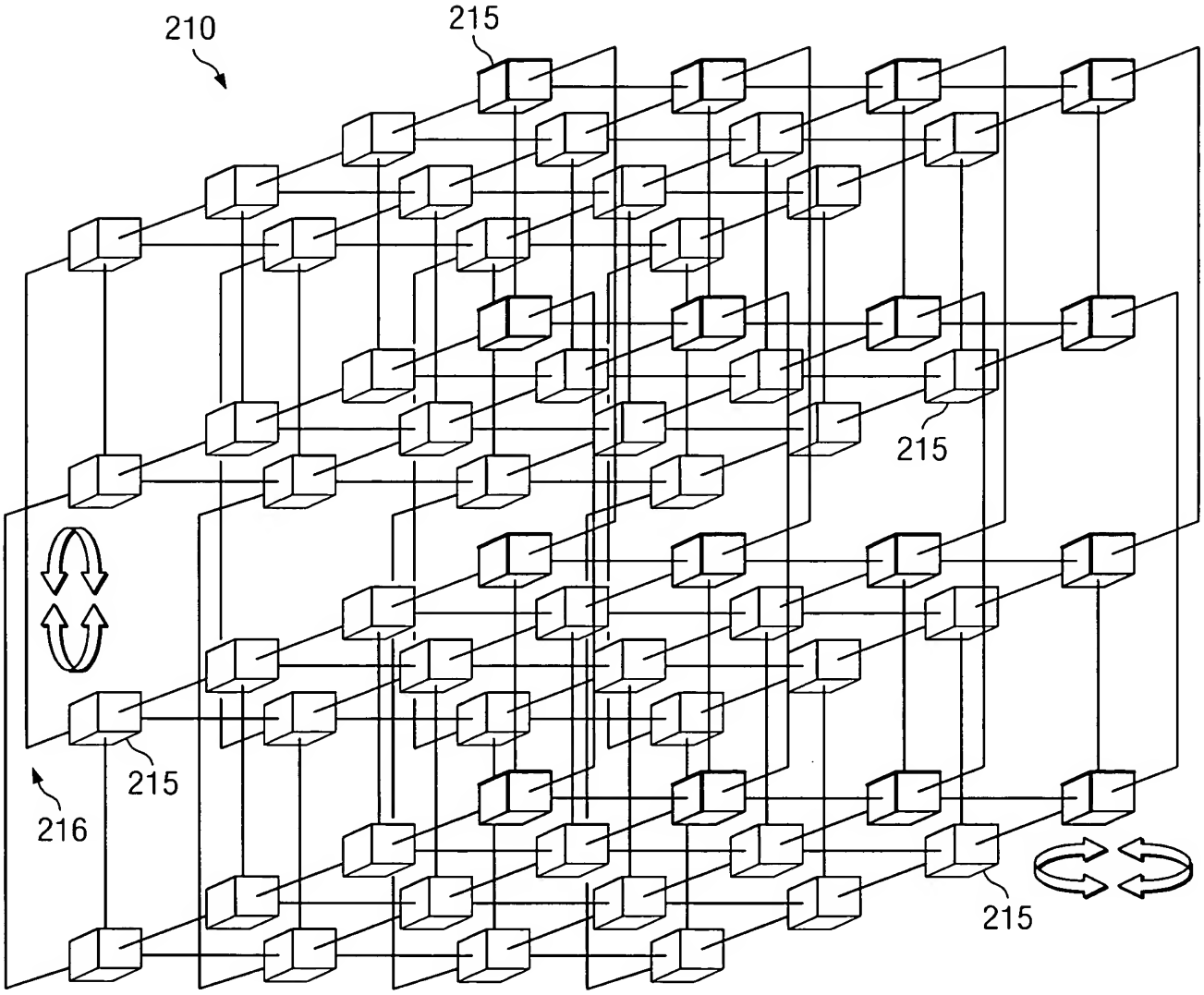


FIG. 2B

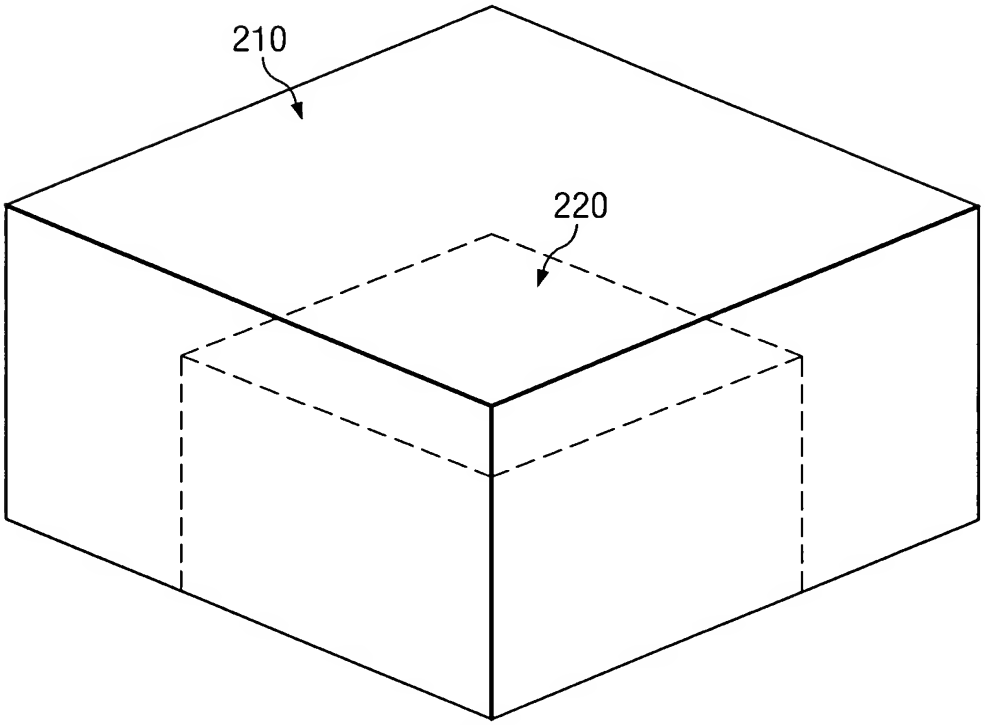


FIG. 2C

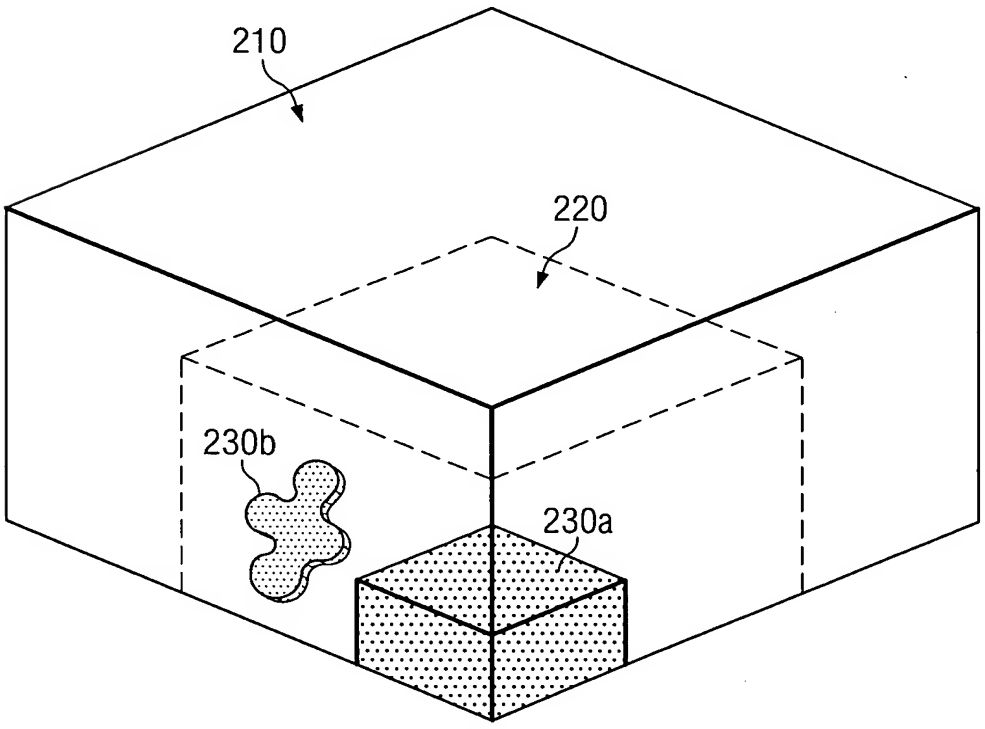


FIG. 2D

FIG. 3A

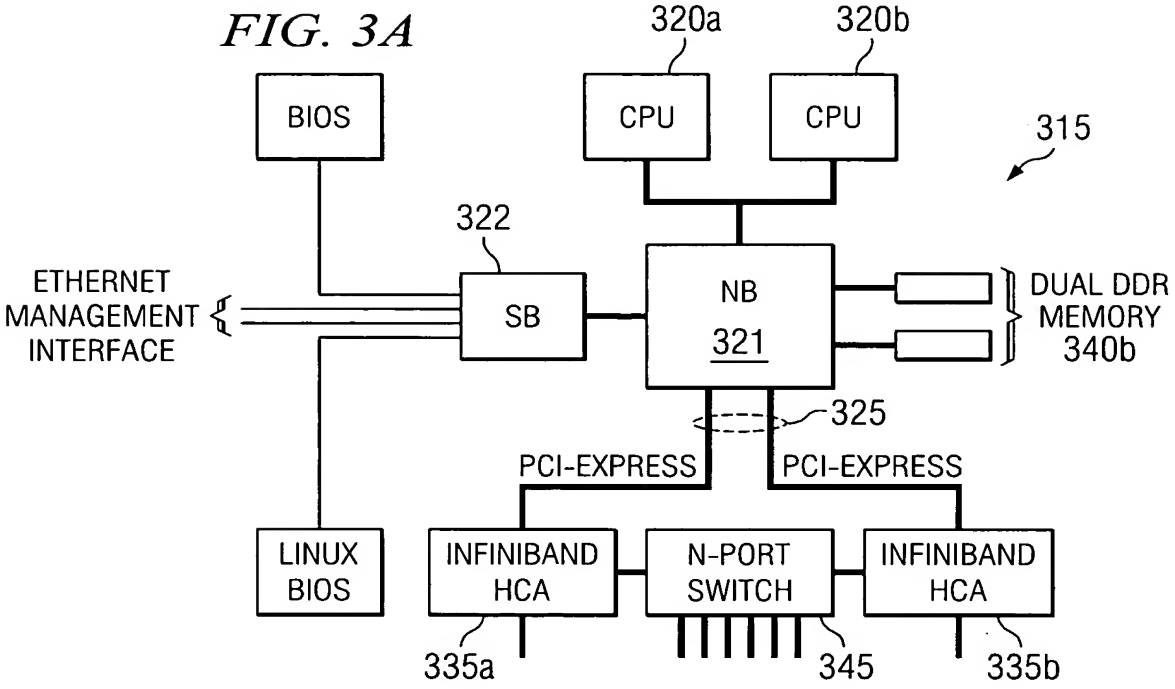

320a
320b
BIOS
CPU
CPU
315
322
ETHERNET
MANAGEMENT
INTERFACE
SB
NB
321
DUAL DDR
MEMORY
340b
325
PCI-EXPRESS
PCI-EXPRESS
LINUX
BIOS
INFINIBAND
HCA
N-PORT
SWITCH
INFINIBAND
HCA
335a
345
335b

FIG. 3B

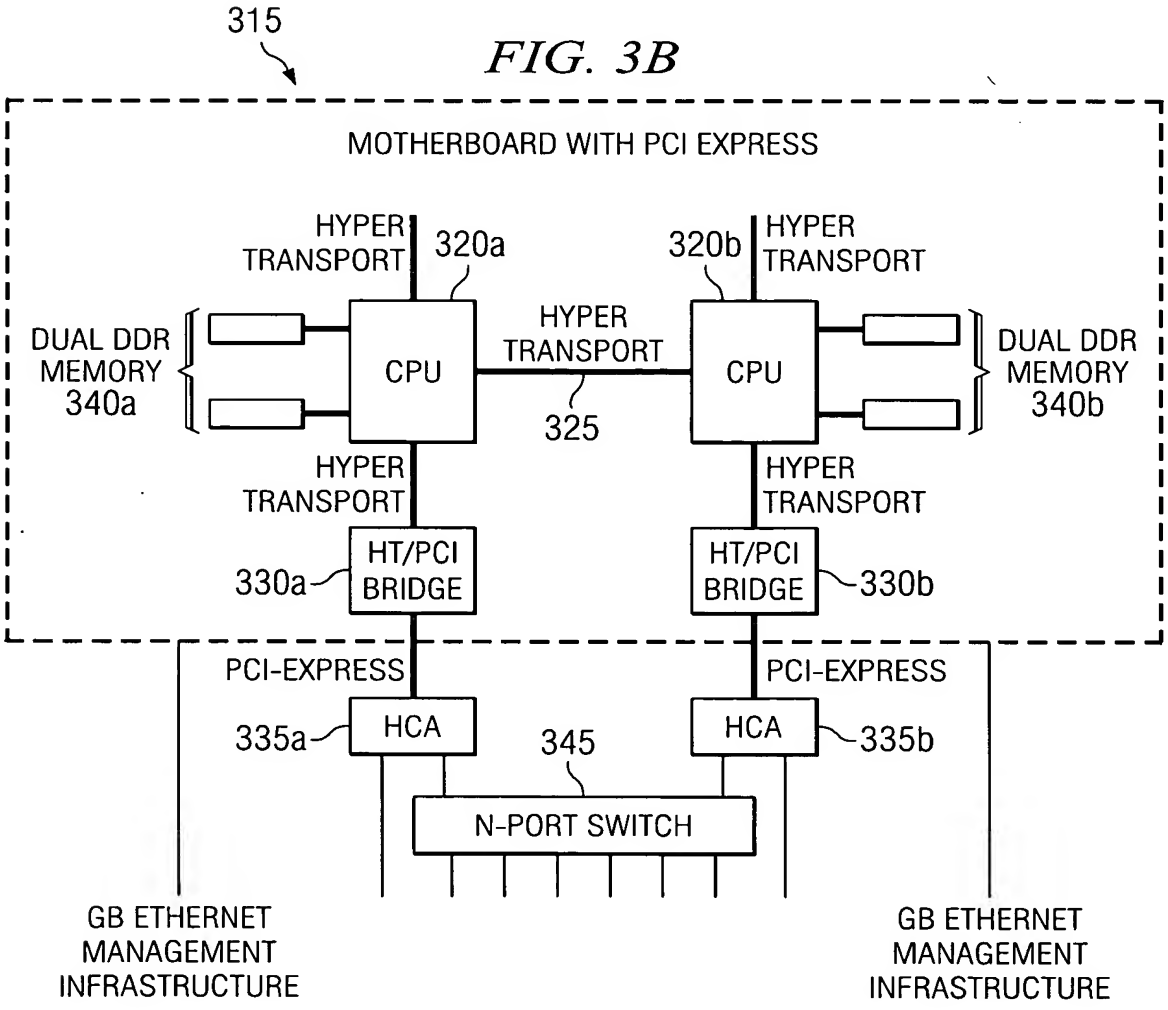
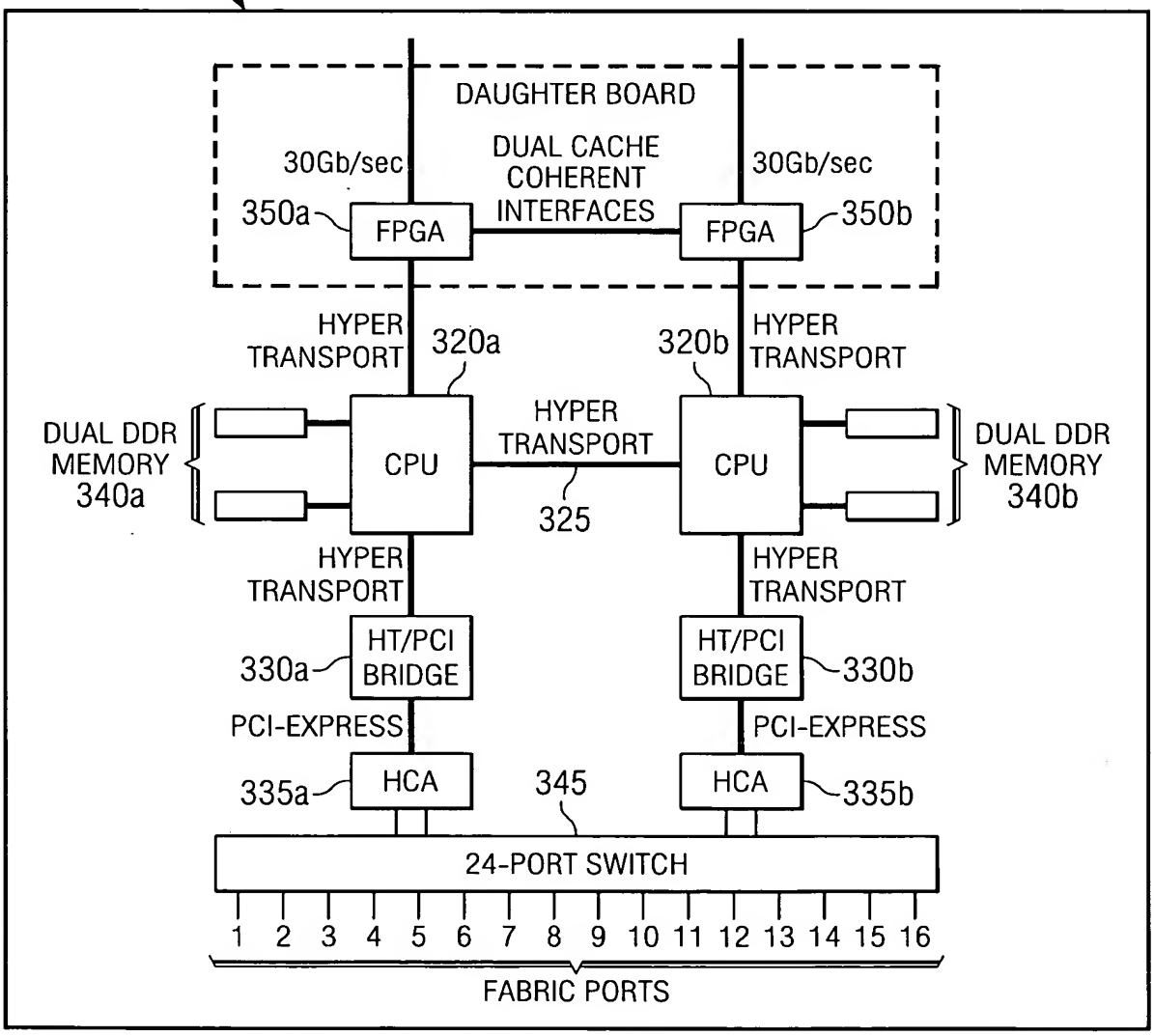

315
MOTHERBOARD WITH PCI EXPRESS
HYPER
TRANSPORT
320a
320b
HYPER
TRANSPORT
DUAL DDR
MEMORY
340a
CPU
HYPER
TRANSPORT
325
CPU
DUAL DDR
MEMORY
340b
HYPER
TRANSPORT
HYPER
TRANSPORT
330a
HT/PCI
BRIDGE
HT/PCI
BRIDGE
330b
PCI-EXPRESS
PCI-EXPRESS
335a
HCA
345
HCA
335b
N-PORT SWITCH
GB ETHERNET
MANAGEMENT
INFRASTRUCTURE
GB ETHERNET
MANAGEMENT
INFRASTRUCTURE

FIG. 3C


315
DAUGHTER BOARD
30Gb/sec
DUAL CACHE COHERENT INTERFACES
30Gb/sec
350a
FPGA
FPGA
350b
HYPER TRANSPORT
320a
320b
HYPER TRANSPORT
DUAL DDR MEMORY 340a
CPU
HYPER TRANSPORT
CPU
DUAL DDR MEMORY 340b
325
HYPER TRANSPORT
HYPER TRANSPORT
330a
HT/PCI BRIDGE
HT/PCI BRIDGE
330b
PCI-EXPRESS
PCI-EXPRESS
335a
HCA
345
HCA
335b
24-PORT SWITCH
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
FABRIC PORTS

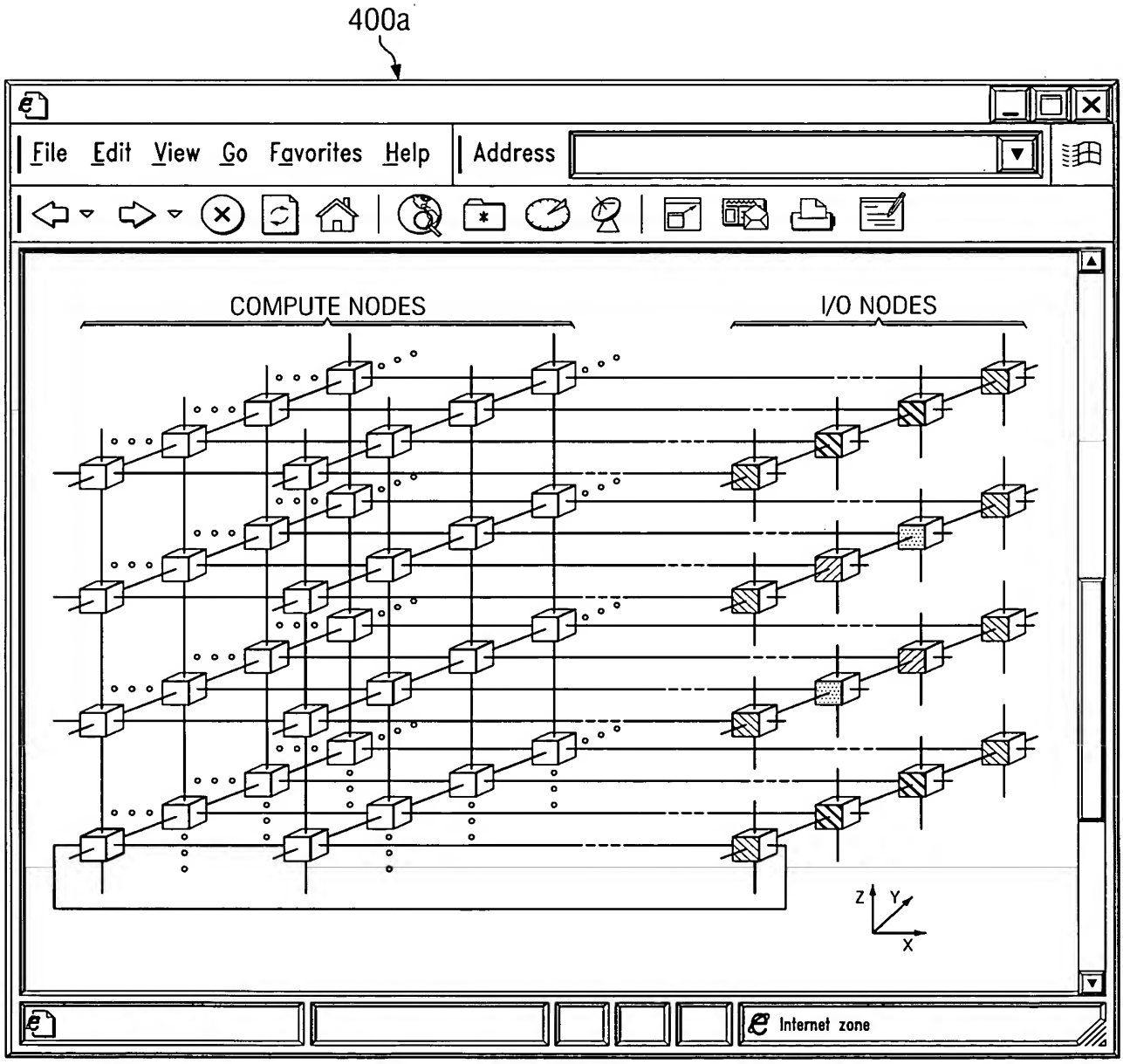

400a
File Edit View Go Favorites Help
Address
COMPUTE NODES
I/O NODES
Z
Y
X
Internet zone

FIG. 4A

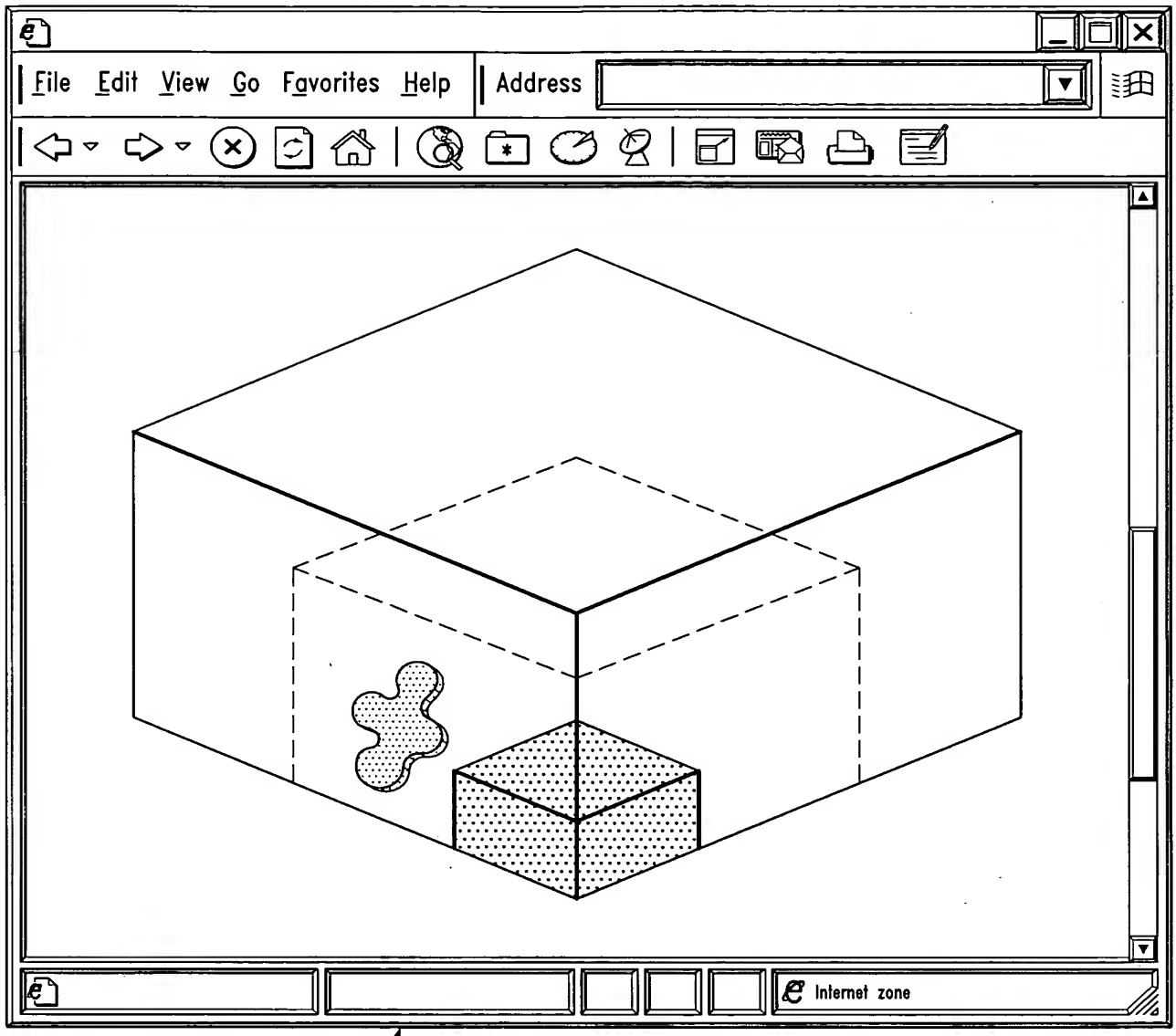

File Edit View Go Favorites Help
Address
Internet zone
400b

FIG. 4B

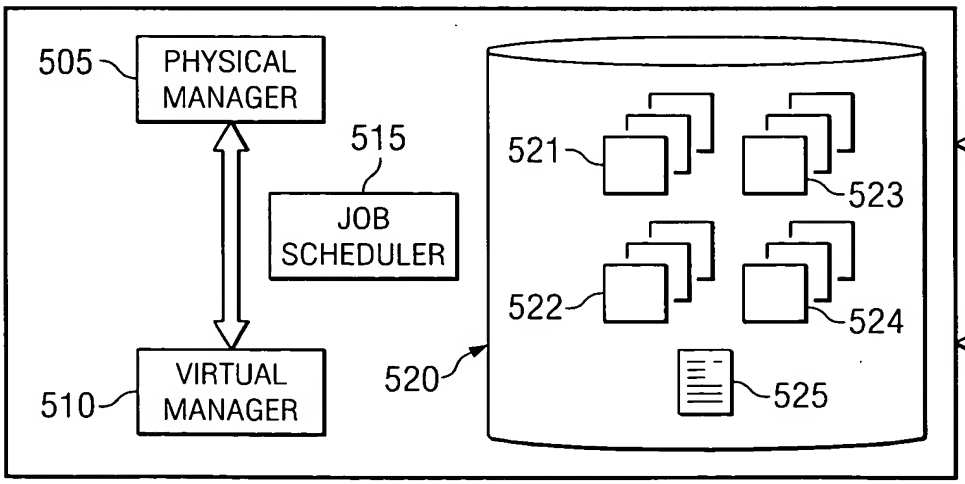
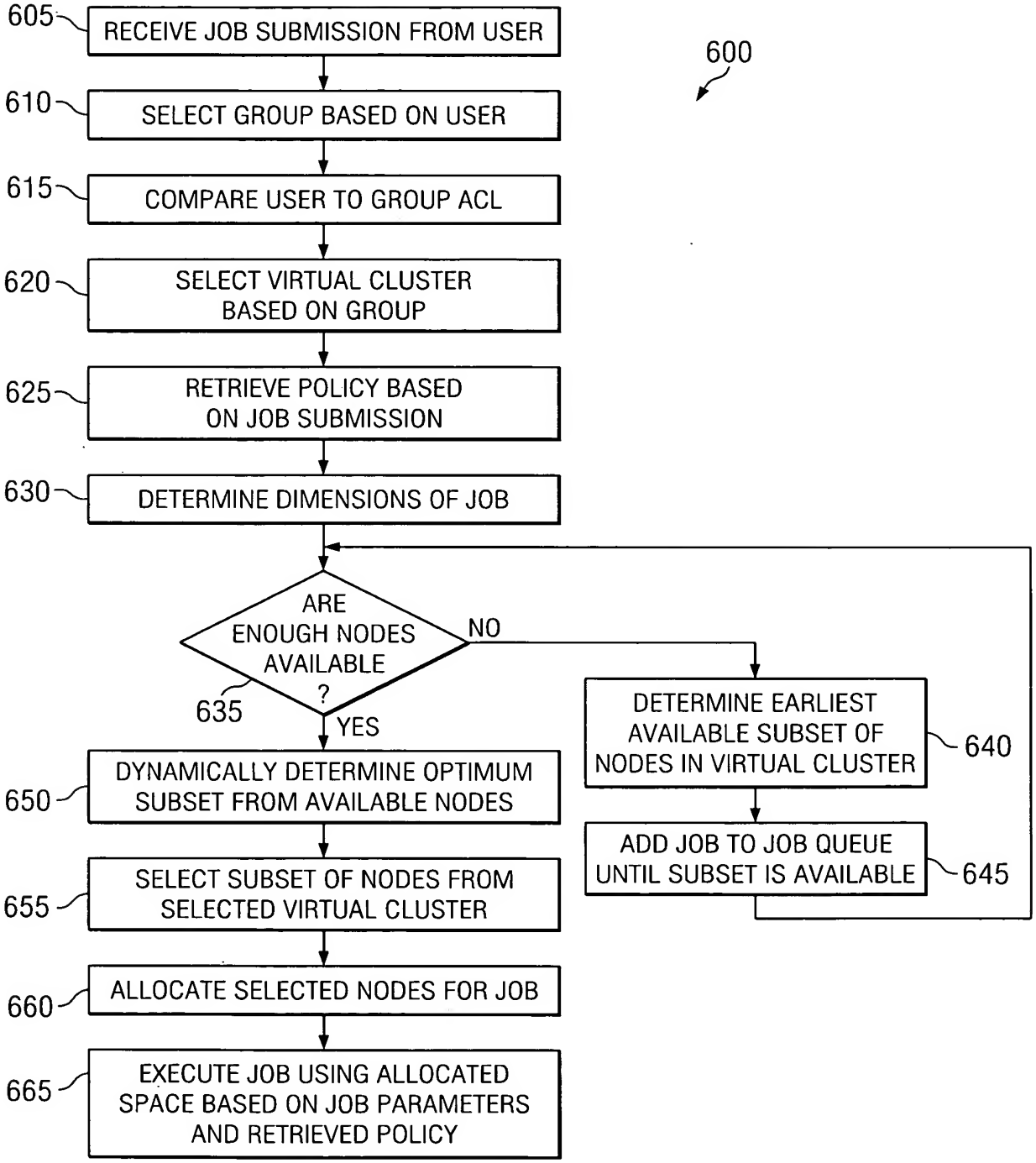

505
PHYSICAL MANAGER
515
JOB SCHEDULER
510
VIRTUAL MANAGER
520
521
523
522
524
525
JOB
POLICIES
500

FIG. 5

FIG. 6


600
605
RECEIVE JOB SUBMISSION FROM USER
610
SELECT GROUP BASED ON USER
615
COMPARE USER TO GROUP ACL
620
SELECT VIRTUAL CLUSTER BASED ON GROUP
625
RETRIEVE POLICY BASED ON JOB SUBMISSION
630
DETERMINE DIMENSIONS OF JOB
ARE ENOUGH NODES AVAILABLE ?
635
NO
YES
DETERMINE EARLIEST AVAILABLE SUBSET OF NODES IN VIRTUAL CLUSTER
640
ADD JOB TO JOB QUEUE UNTIL SUBSET IS AVAILABLE
645
650
DYNAMICALLY DETERMINE OPTIMUM SUBSET FROM AVAILABLE NODES
655
SELECT SUBSET OF NODES FROM SELECTED VIRTUAL CLUSTER
660
ALLOCATE SELECTED NODES FOR JOB
665
EXECUTE JOB USING ALLOCATED SPACE BASED ON JOB PARAMETERS AND RETRIEVED POLICY

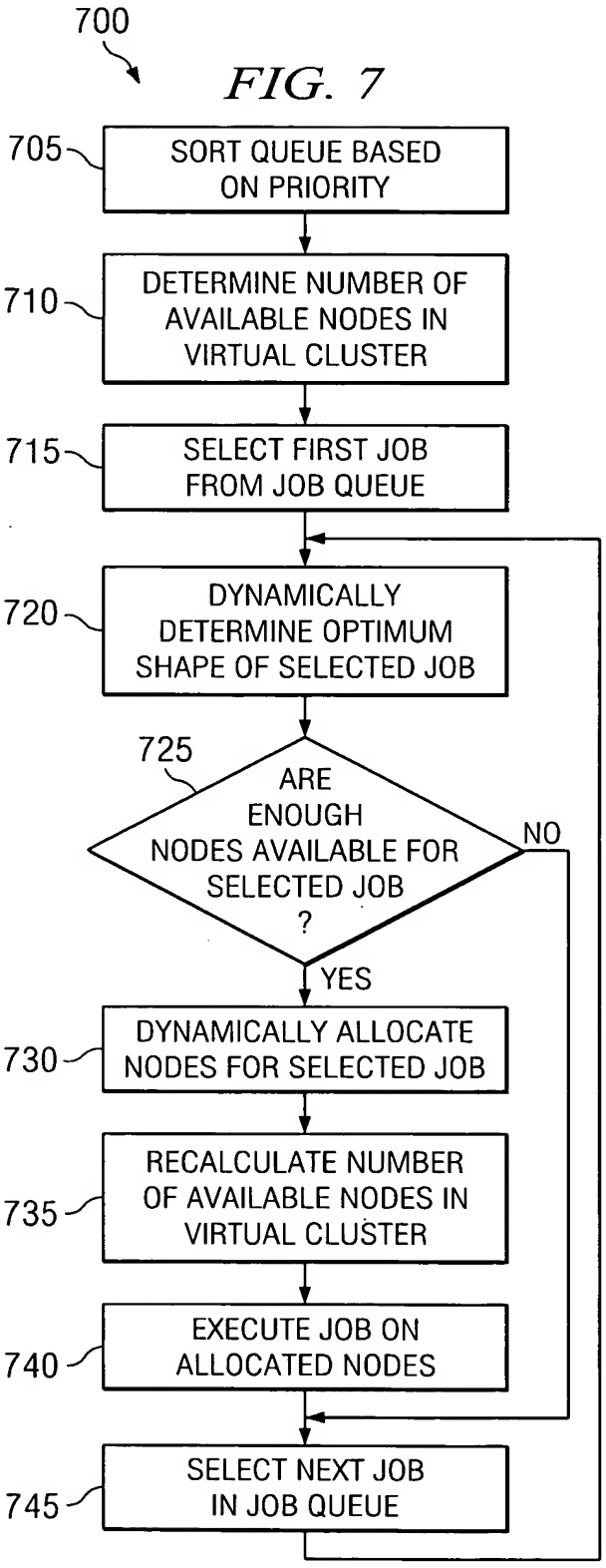

700
705
SORT QUEUE BASED ON PRIORITY
710
DETERMINE NUMBER OF AVAILABLE NODES IN VIRTUAL CLUSTER
715
SELECT FIRST JOB FROM JOB QUEUE
720
DYNAMICALLY DETERMINE OPTIMUM SHAPE OF SELECTED JOB
725
ARE ENOUGH NODES AVAILABLE FOR SELECTED JOB ?
NO
YES
730
DYNAMICALLY ALLOCATE NODES FOR SELECTED JOB
735
RECALCULATE NUMBER OF AVAILABLE NODES IN VIRTUAL CLUSTER
740
EXECUTE JOB ON ALLOCATED NODES
745
SELECT NEXT JOB IN JOB QUEUE

FIG. 7

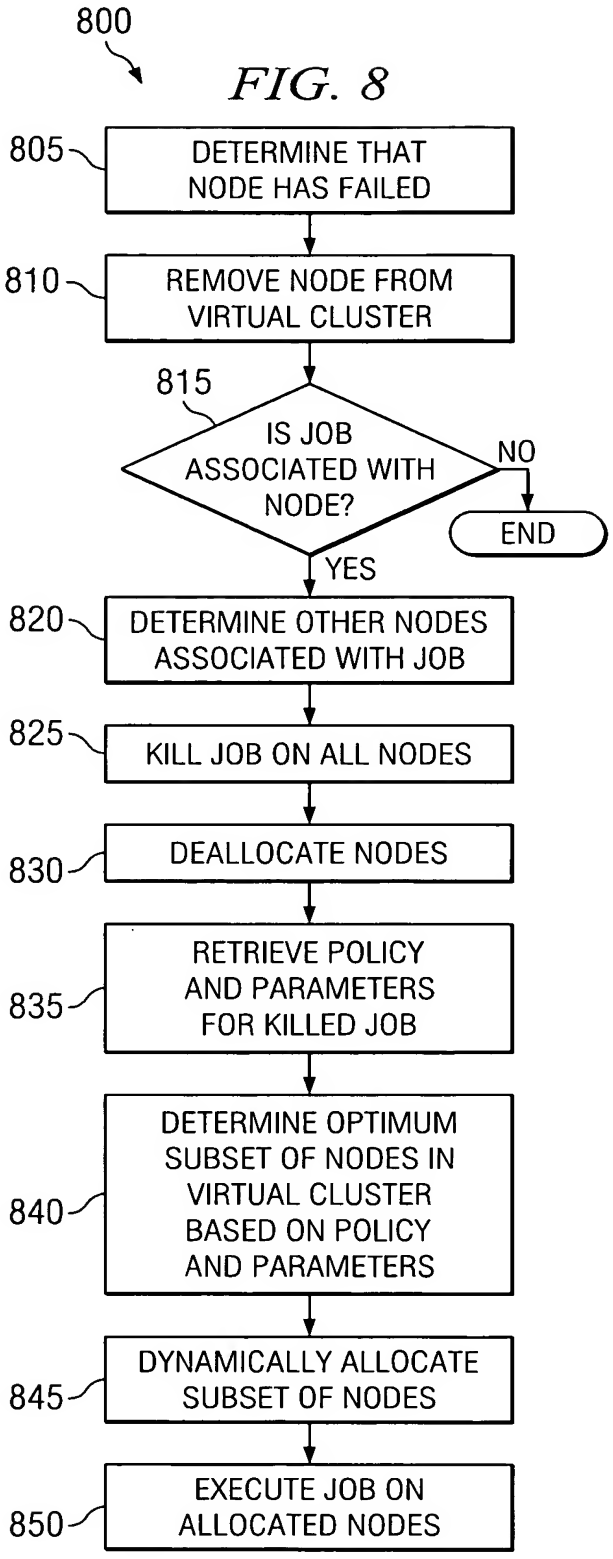

800
805
DETERMINE THAT NODE HAS FAILED
810
REMOVE NODE FROM VIRTUAL CLUSTER
815
IS JOB ASSOCIATED WITH NODE?
NO
END
YES
820
DETERMINE OTHER NODES ASSOCIATED WITH JOB
825
KILL JOB ON ALL NODES
830
DEALLOCATE NODES
835
RETRIEVE POLICY AND PARAMETERS FOR KILLED JOB
840
DETERMINE OPTIMUM SUBSET OF NODES IN VIRTUAL CLUSTER BASED ON POLICY AND PARAMETERS
845
DYNAMICALLY ALLOCATE SUBSET OF NODES
850
EXECUTE JOB ON ALLOCATED NODES

FIG. 8